

Supplementary Note 2

1 Derivation of the fractional DNA occupancy as a function of buffered metal concentration

1.1 Chemical equilibria and mass balance

In the following system:



The constants K_{1-4} are connected by the following relationship

$$K_3 = \frac{K_1}{K_2} K_4.$$

The fraction of DNA bound to sensor protein (θ_D) and the sub-fraction bound solely to metalated sensor protein (θ_{DM}) are defined at each metal concentration as:

$$\theta_D = \frac{[PD] + [PMD]}{[D_T]}, \quad (1)$$

$$\theta_{DM} = \frac{[PMD]}{[D_T]}, \quad (2)$$

where

$$[D_T] = [D] + [PD] + [PMD]$$

$[D_T]$ is the concentration of DNA targets and its value is independent of the buffered metal concentration $[M]$. At any given buffered metal concentration, the total protein concentration $[P_T]$ is:

$$[P_T] = [P] + [PM] + [PD] + [PMD].$$

The protein abundance and fractional DNA occupancy are both dependent on $[M]$. Relating linearly $[P_T]$ and θ_D for co-repressors and de-repressors:

$$\frac{[P_T] - [P_0]}{[P_1] - [P_0]} = \frac{\theta_D - \theta_{D0}}{\theta_{D1} - \theta_{D0}}, \quad (3)$$

where $[P_0]$ and θ_{D0} are the total protein concentration and the fractional DNA occupancy at low cognate metal concentration, respectively, and $[P_1]$ and θ_{D1} are the equivalent values calculated high cognate

metal concentration. An equivalent equation can be written for CueR and ZntR-like activators, with a linear relationship between $[P_T]$ and θ_{DM} :

$$\frac{[P_T] - [P_0]}{[P_1] - [P_0]} = \frac{\theta_{DM} - \theta_{DM0}}{\theta_{DM1} - \theta_{DM0}}. \quad (4)$$

The numerical values of θ_{D0} , θ_{D1} , θ_{DM0} and θ_{DM1} can (in the first instance, see 1.3.2 and 1.3.4) be derived from the computational approach we recently developed¹³.

1.2 Variables

To simplify derivation, variables and constants are renamed as follows:

$$\begin{array}{lll} x = [P], & v = [PD], & a = K_1, \\ y = [M], & w = [PMD], & b = K_2, \\ z = [D], & P_T = [P_T], & c = K_3, \\ D_T = [D_T], & P_0 = [P_0], & d = K_4. \\ u = [PM], & P_1 = [P_1], & \end{array}$$

1.3 Derivation of equations

Here the equations expressing θ_D and θ_{DM} as a function of the buffered metal concentration y are derived. From the chemical equilibria and mass balances the following relationships between variables are obtained:

$$a = \frac{u}{xy}, \quad (5)$$

$$b = \frac{w}{vy}, \quad (6)$$

$$d = \frac{w}{zu}, \quad (7)$$

$$D_T = z + v + w, \quad (8)$$

$$P_T = x + u + v + w = x + u + D_T - z. \quad (9)$$

Using (5), (6) and (7), u and w can be expressed as:

$$u = axy, \quad w = bvy = duz. \quad (10)$$

With (8) v can be determined:

$$D_T = z + v + w = z + v + duz \Rightarrow v = D_T - z - duz.$$

By substituting v into (10) z can be expressed as

$$byv = by(D_T - z - duz) = duz,$$

$$z = \frac{D_T by}{du + by + bduy}.$$

1.3.1 Derivation of equation to determine θ_D for co-repressors and de-repressors

Equation (3) can be rewritten as

$$P_T = \frac{\Delta P}{\Delta \theta_D} \theta_D + A_D, \quad (11)$$

where

$$\begin{aligned} \Delta P &= P_1 - P_0, \\ \Delta \theta_D &= \theta_{D1} - \theta_{D0}, \\ A_D &= P_0 - \frac{\theta_{D0} \Delta P}{\Delta \theta_D}. \end{aligned} \quad (12)$$

The definition of θ_D from (1) can be rearranged as:

$$\theta_D = \frac{v+w}{D_T} = 1 - \frac{z}{D_T}, \quad (13)$$

and substituted into (11), giving:

$$\begin{aligned} P_T &= \frac{\Delta P}{\Delta \theta_D} \theta_D + A_D, \\ x + u + D_T - z &= \frac{\Delta P}{\Delta \theta_D} \left(1 - \frac{z}{D_T} \right) + A_D, \\ x + u - C_1 z &= C_2, \end{aligned} \quad (14)$$

where

$$C_1 = 1 - \frac{\Delta P}{\Delta \theta_D \cdot D_T}, \quad (15)$$

$$C_2 = \frac{\Delta P}{\Delta \theta_D} + A_D - D_T. \quad (16)$$

z and u can be substituted into (14) to find x

$$\begin{aligned} x + axy - C_1 \cdot \frac{D_T by}{adxy + by + abdxy^2} &= C_2, \\ x(1 + ay)(adx + b + abdxy) - C_1 D_T b &= C_2(adx + b + abdxy), \\ x^2(1 + ay)(ad + abdy) + x[b(1 + ay) - C_2(ad + abdy)] - C_1 D_T b - C_2 b &= 0. \end{aligned} \quad (17)$$

Equation (17) expresses the variable x (corresponding to $[P]$) as a function of y , the buffered metal concentration $[M]$. It can be rewritten, introducing the coefficients α, β, γ , as

$$\alpha x^2 + \beta x + \gamma = 0,$$

where

$$\begin{aligned} \alpha &= (1 + ay)(ad + abdy), \\ \beta &= b(1 + ay) - C_2(ad + abdy), \\ \gamma &= -b(C_1 D_T + C_2). \end{aligned}$$

By substituting the numerical values of the constants, it is noticed that, for any given value of y , α is positive and γ is negative. The two solutions of the quadratic equation, x_1 and x_2 , are linked by the relationship

$$x_1 x_2 = \frac{\gamma}{\alpha}$$

As $\frac{\gamma}{\alpha} < 0$, one of the two solutions is negative and hence is meaningless. Therefore the positive solution is retained

$$x = \frac{-\beta + \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}. \quad (18)$$

Finally, fractional DNA occupancy θ_D can be derived as a function of the buffered metal concentration only as follows

$$\begin{aligned} \theta_D &= 1 - \frac{z}{D_T} \\ \theta_D &= 1 - \frac{by}{du + by + bduy} \\ \theta_D &= 1 - \frac{b}{b + (ad + abdy)x}. \end{aligned} \quad (19)$$

Equation (19) can be solved using an electronic spreadsheet to calculate θ_D given a range of buffered metal concentrations (y), considering that the relationship between x and y is given by (18) (Supplementary Dataset).

1.3.2 Calculation of θ_{D0} and θ_{D1} (an alternative approach)

A simplified form of equation (19) can be used to calculate numerical values of θ_{D0} and θ_{D1} (3) from an electronic spreadsheet (Supplementary Dataset).

1.3.3 Derivation of equation to determine θ_{DM}

Equation (4) can then be rewritten as

$$P_T = \frac{\Delta P}{\Delta \theta_{DM}} \theta_{DM} + A_{DM}, \quad (20)$$

where

$$\begin{aligned} \Delta \theta_{DM} &= \theta_{DM1} - \theta_{DM0}, \\ A_{DM} &= P_0 - \frac{\theta_{DM0} \Delta P}{\Delta \theta_{DM}}. \end{aligned}$$

and then using (9) and (2)

$$x + u + D_T - z = B \frac{w}{D_T} + A_{DM},$$

where

$$B = \frac{\Delta P}{\Delta \theta_{DM}}. \quad (21)$$

w , z and u can be substituted to find x

$$\begin{aligned} x + axy - \frac{bD_T}{adx + b + abdx} - \frac{B \cdot abdx}{adx + b + abdx} &= A_{DM} - D_T, \\ x^2(1 + ay)(ad + abdy) + x[b(1 + ay) - Babdy - (A_{DM} - D_T)(ad + abdy)] - bA_{DM} &= 0. \end{aligned} \quad (22)$$

Similarly to the previous case, equation (22) can be rewritten as

$$\lambda x^2 + \mu x + \nu = 0,$$

where

$$\begin{aligned}\lambda &= (1 + ay)(ad + abdy), \\ \mu &= b(1 + ay) - Babdy - (A_{DM} - D_T)(ad + abdy), \\ \nu &= -bA_{DM}.\end{aligned}$$

Also in this case, $\lambda > 0$ and $\nu < 0$ for any given value of y , the positive solution of the quadratic equation is

$$x = \frac{-\mu + \sqrt{\mu^2 - 4\lambda\nu}}{2\lambda}. \quad (23)$$

The expression of θ_{DM} as a function of y , the buffered metal concentration, is then

$$\begin{aligned}\theta_{DM} &= \frac{w}{D_T} \\ \theta_{DM} &= \frac{duz}{D_T} \\ \theta_{DM} &= \frac{abdx y}{b + (ad + abdy)x}.\end{aligned} \quad (24)$$

Equation (24) can be solved using an electronic spreadsheet to calculate θ_{DM} given a $[M]$ (y) range, considering that the relationship between x and y is given by equation (23) (Supplementary Dataset).

1.3.4 Calculation of θ_{DM0} and θ_{DM1} (an alternative approach)

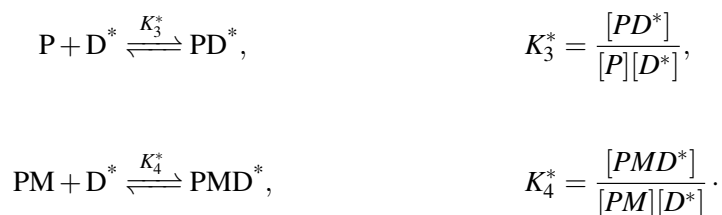
A simplified form of equation (24) can be used to calculate numerical values of θ_{DM0} and θ_{DM1} (4) from an electronic spreadsheet (Supplementary Dataset).

2 Calculation of fractional DNA occupancy at different salt concentrations

DNA affinities have a log-log dependence on salt concentration^{29,30}. Apo-Zur, Zn(II)-Zur and Ni(II)-NikR DNA affinities were experimentally determined at various salt concentrations, and the mean of the regression lines of $\log K_{DNA}$ vs. $\log[\text{salt}]$ plots was used to calculate DNA affinities for the other sensors at 500 mM salt from the values in Table 1 measured at 300 mM. With the K_3 or K_4 values at 500 mM salt, fractional DNA occupancies θ_D and θ_{DM} were calculated using equations (19) and (24).

3 Derivation of fractional DNA occupancy as a function of buffered metal concentration considering sensor binding to non-specific DNA

An excess of non-specific DNA competes *in vivo* with the specific consensus sequences for sensor binding. To incorporate non-specific DNA, here represented as D^* , in the model it is necessary to introduce two additional reactions to the system presented in section 1.1:



The total concentration of non-specific DNA binding sites, $[D_T^*]$, can be calculated by dividing the concentration of available non-specific DNA base pairs (10^{-4} M bp in *E. coli*²⁶) by the average length of DNA binding sequences used as specific targets (33 bp). The mass balance for non-specific DNA is

$$[D_T^*] = [D^*] + [PD^*] + [PMD^*],$$

while the mass balance for the protein, incorporating sensor binding to non-specific DNA, is now

$$[P_T] = [P] + [PM] + [PD] + [PMD] + [PD^*] + [PMD^*].$$

The affinities of Zur for non-specific DNA, K_3^* and K_4^* , were experimentally measured on the *nixA* promoter at 100 mM salt. Non-specific DNA affinities at 300 mM salt were calculated from the slope of the regression line of the Zur $\log K_{DNA}$ vs $\log[\text{salt}]$ and the (small) coupling free energy on non-specific DNA ($\Delta G_c^* = -RT \ln(\frac{K_4^*}{K_3^*})$) was determined for Zur. The non-specific DNA affinities of the other sensors were estimated by maintaining the same proportion between ΔG_c^* and ΔG_c and with K_3^* and K_4^* flanking K_3 for co-repressors and flanking K_4 for the de-repressor and activators (K_3 or K_4 respectively defining the midpoint on a logarithmic scale between K_3^* and K_4^*).

The fraction of specific DNA sites bound to sensor protein (θ_D) and the sub-fraction bound solely to metalated sensor protein (θ_{DM}) are defined at each metal concentration by equations (1) and (2). For co-repressors and de-repressors the total protein concentration $[P_T]$ relates linearly to θ_D as expressed in equation (3), for CueR and ZntR-like activators the relationship between $[P_T]$ and θ_{DM} is given in equation (4).

3.1 Variables

In addition to the variables presented in 1.2, the additional variables and constants are renamed as follows to simplify computation:

$$\begin{aligned} e &= K_3^*, & h &= [D^*] \\ f &= K_4^*, & i &= [PD^*] \\ g &= K_2^*, & j &= [PMD^*] \\ D_T^* &= [D_T^*]. \end{aligned}$$

3.2 Derivation of equations

From the chemical equilibria and mass balances the following relationships are derived:

$$e = \frac{i}{xh}, \tag{25}$$

$$f = \frac{j}{uh}, \tag{26}$$

$$D_T^* = h + i + j, \tag{27}$$

$$P_T = x + u + v + w + i + j = x + u + D_T^* - z + D_T^* - h. \tag{28}$$

Using (25) and (26), i and j can be expressed as

$$i = exh, \quad j = fuh \quad (29)$$

and h can then be derived from (27)

$$h = \frac{D_T^*}{1 + ex + fu}$$

3.2.1 Derivation of equation to determine θ_D for co-repressors and de-repressors

Equation (11) can be rewritten using (13) and (28) as

$$P_T = \frac{\Delta P}{\Delta \theta_D} \theta_D + A_D,$$

$$x + u + D_T - z + D_T^* - h = \frac{\Delta P}{\Delta \theta_D} \left(1 - \frac{z}{D_T}\right) + A_D,$$

$$x + u - C_1 z - h = C_2^*, \quad (30)$$

where

$$C_1 = 1 - \frac{\Delta P}{\Delta \theta_D \cdot D_T},$$

$$C_2^* = \frac{\Delta P}{\Delta \theta_D} + A_D - D_T - D_T^*.$$

The variables u , z and h can be substituted into (30) to find x

$$x + axy - C_1 \cdot \frac{D_T by}{adxy + by + abdxy^2} - \frac{D_T^*}{1 + ex + afxy} = C_2^*,$$

$$\begin{aligned} & x^3(1 + ay)(ad + abdy)(e + afy) + x^2 \left\{ (1 + ay)(ad + abdy) + [b(1 + ay) - C_2^*(ad + abdy)](e + afy) \right\} + \\ & + x \left[b(1 + ay) - C_2^*(ad + abdy) - b(C_1 D_T + C_2^*)(e + afy) - D_T^*(ad + abdy) \right] - b(C_1 D_T + C_2^* + D_T^*) \end{aligned} \quad (31)$$

Equation (31) expresses the variable x (corresponding to $[P]$) as a function of y , the buffered metal concentration $[M]$. It is possible to numerically solve the cubic equation using an electronic spreadsheet or a computing environment. In all the cases examined a positive and two negative solutions were obtained and the positive solution was retained to calculate θ_D from equation (19). Due to the complex analysis required to solve the cubic equation, a Supplementary Dataset is not provided.

3.2.2 Calculation of θ_{D0} and θ_{D1}

A simplified form of equation (31) can be used to calculate numerical values of θ_{D0} and θ_{D1} (3) from an electronic spreadsheet or a computing environment.

3.2.3 Derivation of equation to determine θ_{DM}

Equation (20) can be rewritten using (2), (10) and (28) as

$$x + u + D_T - z + D_T^* - h = B \frac{w}{D_T} + A_{DM}, \quad (32)$$

where is expressed by equation (21). The variables u , z and h can be substituted into (32) to find x

$$x + axy + D_T - \frac{D_T by}{adxy + by + abdxy^2} + D_T^* - \frac{D_T^*}{1 + ex + afxy} = B \frac{abdxy^2}{adxy + by + abdxy^2} + A_{DM}$$

$$x^3(1 + ay)(ad + abdy)(e + afy) + x^2 \left\{ (1 + ay)(ad + abdy) + [b(1 + ay) - Babdy - (A_{DM} - D_T - D_T^*)(ad + abdy)](e + afy) \right\} + x \left[b(1 + ay) - Babdy - (A_{DM} - D_T - D_T^*)(ad + abdy) - b(A_{DM} - D_T^*)(e + afy) - D_T^*(ad + abdy) \right] - bA_{DM} = 0 \quad (33)$$

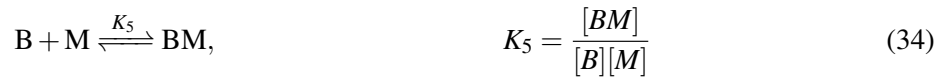
Equation (33) can be numerically solved using an electronic spreadsheet or a computing environment. In all the cases examined a positive and two negative solutions were obtained and the positive solution was used to calculate θ_{DM} from equation (24). Due to the complex analysis required to solve the cubic equation, a Supplementary Dataset is not provided.

3.2.4 Calculation of θ_{DM0} and θ_{DM1}

A simplified form of equation (33) can be used to calculate numerical values of θ_{DM0} and θ_{DM1} (3) from an electronic spreadsheet or a computing environment.

4 Derivation of the relationship between total metal bound to buffer and buffered metal concentration

The buffered metal system is described as



The concentration of buffering species, $[B_T]$, and the total metal concentration in the buffer system, $[M_T]$, are:

$$[M_T] = [M] + [BM] \quad (35)$$

$$[B_T] = [B] + [BM] \quad (36)$$

From equations (35) and (36), $[BM]$ and $[B]$ can be derived and substituted into equation (34):

$$[BM] = [M_T] - [M]$$

$$[B] = [B_T] - [BM] = [B_T] - [M_T] + [M]$$

$$K_5 = \frac{[M_T] - [M]}{([B_T] - [M_T] + [M])[M]}. \quad (37)$$

Equation (37) can be rearranged to express $[M_T]$ as a function of $[M]$

$$[M_T] = \frac{K_5[M]^2 + (K_5[B_T] + 1)[M]}{1 + K_5[M]}. \quad (38)$$

An example of the use of this relationship is shown in Figure 4b.

References

- 29 Record, M. T., Jr., Ha, J. H. & Fisher, M. A. Analysis of equilibrium and kinetic measurements to determine thermodynamic origins of stability and specificity and mechanism of formation of site-specific complexes between proteins and helical DNA. *Methods in Enzymol.* **208**, 291-343 (1991).
- 30 Campanello, G. C. et al. Allosteric inhibition of a zinc-sensing transcriptional repressor: insights into the arsenic repressor (ArsR) family. *J. Mol. Biol.* **425**, 1143-1157 (2013).